AFOSR-TR- 80 - 0283

EXTENSION DIVISION

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE
GRADUATE PROGRAM IN NORTHERN VIRGINIA

P. O. Box 17186
Washington, D. C. 20041
(703) 471-4600

Reliable Software Through Very High-Level Verification

INTERIM SCIENTIFIC REPORT

Grant AFOSR-79-0021

LEVEL

DTIC
SELECTED
APR 24 1980

Richard J. Orgass

February 15, 1980

Technical Memorandum No. 80-2

Submitted to

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Building 410
Bolling Air Force Base
Washington, D.C. 20332

80 4 21 014

public release;
limited.

Located at Dulles International Airport—400 West Service Road

ADA083611

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 80 -0283	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) RELIABLE SOFTWARE THROUGH VERY HIGH-LEVEL VERIFICATION		5. TYPE OF REPORT & PERIOD COVERED Interim rept's
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard J. Orgass		8. CONTRACT OR GRANT NUMBER(s) AFOSR-79-0021
9. PERFORMING ORGANIZATION NAME AND ADDRESS Va. Poly. Inst. & St. University Department of Computer Science Washington, DC 20041		10. PROGRAM ELEMENT. PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A2
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE 15 February 1980
		13. NUMBER OF PAGES 15
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

VPI-TM-80-3,

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The work of the research period that is the subject of this report is best described as preparation for the substantial work of the grant. When this work began, a substantial body of software to be used for the research was in operation on a DEC-10 and this software was transported to the Virginia Tech VM/CMS system. Significant difficulties were encountered in bringing this software into satisfactory operation in this environment.

TABLE OF CONTENTS

1. Overview

The work of the research period that is the subject of this report is best described as preparation for the substantial work of the grant. When this work began, a substantial body of software to be used for the research was in operation on a DEC-10 and this software was transported to the Virginia Tech VM/CMS system. Significant difficulties were encountered in bringing this software into satisfactory operation in this environment.

In spite of the fact that SIMULA implementations are closely controlled by the Norwegian Computing Center and that the semantics of the language is expected to be the same in all implementations, different implementations have varying numbers of system errors. It turned out to be the case that the programs used in this research exposed many problems in the IBM SIMULA implementation.

The VM/CMS system is nominally an interactive system but it lacks many of the capabilities that one expects in a modern interactive computing system. Since these capabilities are essential for the work of the grant, it was necessary to devote considerable effort to creating a suitable working environment in VM/CMS.

The essential software of the grant is now operational in VM/CMS and is available for the substantial work that is the principal objective of this work. In addition, tools for building more advanced software are available and are documented adequately for the work of the grant.

The specific pieces of software that have been constructed are as follows:

(1) An interactive incremental verifier has been brought into reliable and convenient operation on VM/CMS.

(2) A partial implementation of APL based on the formal semantics of APL being developed under this grant has been brought into reliable operation.

(3) An interactive file system with high quality terminal interfaces has been designed and partly implemented. Extensions to this system have been designed and they will be implemented as needed.

(4) A set of programs for parsing arbitrary SLR(1) languages and for manipulating the parse trees of such languages has been constructed. The implemented facilities include pattern matching on parse trees.

Since substantial software development was required, less
work than anticipated was devoted to the more abstract parts of
the research. However, there is one major accomplishment to
report in this area. A clean proof of the correctness of data
structures in the implementation of APL was constructed. With
this proof in hand, the proof of the remainder of the implementa-
tion should be quite straightforward. Approximately one third of
a research monograph on this work has been completed and the
existing manuscript will be submitted for publication review in
the near future.

Until the end of October, this research work also suffered
from the lack of reliable data communications between the North-
ern Virginia Graduate Center and the Virginia Tech campus in
Blacksburg. The Commonwealth telephone system, SCATS, was used
for data communications. C&P Telephone (provider of this ser-
vice) was unable to repair SCATS so that the network would meet
the specifications. Therefore, at the end of October 1979, C&P
provided foreign exchange lines from Blacksburg to the Northern
Virginia Graduate Center at it's expense. In December 1979, Vir-
ginia Tech installed a high quality data communications system
between Blacksburg and Northern Virginia. Once the data communi-
cations problems were solved, it was possible to work far more
effectively: the time that had been wasted on transmission error
correction and disconnects was devoted to useful work. Produc-
tivity has substantially increased.

The software development that was completed as well as prob-
lems that were encountered are summarized in this report.
Detailed documentation of much of this work has been published as
Technical Memoranda. A list of these reports appears at the end
of this document and the reports themselves are attached to the
original of this report.

2. APL Semantics and Implementation

Substantial effort was required to install the partial
implementation of APL that was available when this grant began on
the Virginia Tech CMS system. A summary of some of the problems
follows. After this discussion, there is a brief summary of the
status of the work on the semantics of APL and the implementation
of APL.

An unreasonable amount of time was devoted to tracking down
apparent program errors that were found to be a consequence of
the fact that certain characters (e.g., {, }, ~, |,]) have more
than one EBCDIC code. Different system utilities provide differ-
ent translations from ASCII to EBCDIC. All of the known charac-
ter ambiguities can now be corrected by a program that was
created for this purpose.

During the research period, five different SIMULA compilers were used. At the beginning of the research period, MVS SIMULA, Version 6.00 was available at Virginia Tech. The partial implementation of APL exhibited a major error (as well as minor errors) in this compiler/run time system. When the problem was reported to the Norwegian Computing Center (SIMULA supplier), Version 6.02 of MVS SIMULA was provided to repair the problem.

This system made it possible to execute much of the interpreter but the terminal dialog was unacceptable as an approximation to the APL terminal interface. At this point, it appeared that a substantial assembly coding effort would be required to reasonably approximate the APL terminal interface. Just as a decision to write this assembly code was to be made, a version of IBM SIMULA specifically designed for CMS was announced by Imperial College of Science and Technology. This version has a few windows into the operating system and it appeared to be possible to provide a reasonable terminal interface with a minimum amount of assembly coding. Therefore, this system was acquired and installed at Virginia Tech. While this SIMULA system solved some problems, other system errors were present in this version.

The Principal Investigator was engaged in a regular correspondence with the Norwegian Computing Center during this period and a number of his suggestions for enhancements to IBM SIMULA were included in Version 7.00 of MVS SIMULA and this system was installed when Version 6.02 expired. This new system repaired some of the problems with earlier versions but there was another internal error that was corrected by the Norwegian Computing Center.

A significant fragment of the actual code for the APL implementation appears in the manuscript of a monograph and, therefore, a suitable publication format of the program text is required. IBM SIMULA systems before Version 7.00 were upper case only and to provide publication format program text a program to format SIMULA programs was imported from DEC-10 SIMULA and modified for use in the CMS environment. Until August 1979 two copies of program text were maintained: an executable copy and a publication copy. When Version 7.00 of IBM SIMULA became available, the execution text was destroyed because the publication text serves both purposes. This enhancement of the IBM SIMULA system substantially reduced the overhead associated working with SIMULA for this purpose; the readability of programs was also significantly enhanced.

The work on file system development summarized in Section 4 provides both the terminal interface and other support for the APL implementation and is an essential part of the work on the APL implementaion.

The extensive programming effort that was required to bring both the APL implementation and the interactive verifier dis-

cussed in Section 3 into adequate operation in the CMS environment precluded a significant amount of work on the formal semantics of APL and the verification of the implementation. However, some progress was made on this work.

A clean proof of the correctness of the implementation of APL individuals with respect to the formal semantics has been completed. This part of the proof of correctness proved to be quite delicate and a number of false starts were made before a clean proof was completed. While working on the final version of this proof, a number of changes in the program which have increased efficiency as well as clarity were made. The form of the results on the correctness of the implementation of APL individuals will substantially simplify the proof of correctness of the remainder of the implementation.

3. Verification of Programs

Bringing the interactive verifier that was implemented in DEC-10 SIMULA into reliable and useable operation in IBM SIMULA was a long and difficult process and adequately reliable operation was first acheived in January 1980. The difficulties described above for the APL implementation caused more serious problems when working with the verifier.

A genuinely interactive verifier makes heavy demands on the file management capabilities of an operating system and it was quickly discovered that an enhanced file system would be needed to provide adequate facilities while working with the verifier. The work on a file system described in Section 4 provides essential support for the verifier in its present form and some of the capabilities that are not yet being used will be essential as the program is extended to verify APL programs.

The main problem that was encountered with the verifier is related to a number of serious errors in the SIMULA compiler/run time system. This program makes very heavy demands on the storage manager, the procedure calling mechanism and the coroutine linkage mechanism and the implementation of these aspects of SIMULA is particularly weak in IBM SIMULA. The Norwegian Computing Center has devoted substantial efforts to solving these problems and each version of IBM SIMULA has new code for this part of the system. At each step in this improvement, small additional problems and incompatibilities arise.

Each version change of the SIMULA system required careful checking of the program and subtle modifications to avoid the current version of bugs in the system. It has been observed that code which was introduced to cause the program to execute more reliably in one version caused problems in a later version.

Version 7.00 of MVS SIMULA is much more reliable than earlier versions but the run time system caused unpredictable terminations for protection or addressing exceptions. A fortunate combination of events led to a useful solution of these problems.

The University Computing Center changed from VM/CMS 5.8 to VM/CMS 6.5 and at the same time Version 7.00 of CMS SIMULA became available. With the active support of Computing Center staff members, it was determined that most of the run time system problems can be avoided by using the CMS SIMULA compiler with the MVS SIMULA run time system under VM/CMS 6.5. Although this combination is not without problems, it appears to be adequately reliable for most purposes.

Since the verifier came into reliable operation, the program has been divided into several logically related segments of code and this code now forms the basis for a laboratory to construct a variety of programs that are useful both in this research and in other research and instructional activities.

Having achieved reliable operation of the verifier, rapid progress is being made at converting the verifier into a verification system that is independent of the syntax and semantics of the language in which programs to be verified are written. This enhancement is being completed by modifying the verifier so that it is completely driven by inputs read from files. It appears to be the case that this can be accomplished without increasing the overhead associated with running the verifier. [There is, of course, the small cost associated with reading data from a file instead of having the data in the program; this is negligible.]

A report of known problems in Version 7.00 of IBM SIMULA has been transmitted to the Norwegian Computing Center and repairs are expected.

Work on the verifier consumed a significant fraction of the human and computing resources of the grant. It is, however, possible to report that a reasonably reliable program with a good user interface is available for the important research work of this grant.

4. File System Development

Since the principal objective of this research program is the verification of APL programs, work on the design and implementation of an input/output system for SIMULA requires some explanation.

Both the incremental verifier and the APL implementation which are an important part of this research are strongly inter-

active programs and these programs require dynamic file naming (that is, the ability to name and open a file during execution) if they are to function in an acceptable manner. The CMS SIMULA system as provided by the Norwegin Computing Center uses OS simulation macros to perform input/output. This system suffers from two important limitations when working with interactive programs:

The CMS simulation of OS input/output requires that file definitions be provided to a program before execution begins. Moreover, once a file definition has been executed, the file is known to a program by a DD name rather than the file name. When working with both the verifier and the APL implementation, one does not in general know which files will be read or written during execution. One possible solution to this problem is to issue a large number of file defintions before program execution and then retain a written record of the correspondence between DD names and file names. Even using the EXEC processor of CMS, this was found to be extremely irritating and the source of many errors.

Terminal support for program input/output in the IBM SIMULA implementation is a byproduct of this use of OS simulation macros. Output written to the terminal is double buffered and, therefore, extensive program modifications were needed to write extra blank lines to make sure that a prompt is printed before the response is read from the terminal. An empty line of input is treated as an end-of-file and an accidental hitting of the return key causes the termination of execution with a loss of all work.

In early work with CMS, ad hoc solutions to these problems were employed and it became quite clear that these problems would persist and become more and more difficult. Therefore, a comprehensive solution to these problems was undertaken.

The first approach to the solution of this problem was the design and implementation of a SIMULA class DIALOG which contains procedures to provide dynamic file naming and a minimal set of primitives for terminal dialog. In addition, a number of procedures which are available in the DEC-10 SIMULA library but which are unavailable in the IBM SIMULA library were added to this class.

This first version of DIALOG greatly simplified work with the verifier but it quickly became obvious that there were still too many possibilities for errors. In addition, it was still impossible to reproduce the prompt by indentation that is used in APL as well as to provide some prompts from the verifier in a reasonable way.

A second verison of DIALOG was constructed to provide additional security and prevent unexpected termination of execution because of some trivial error. In addition, a small amount of

assembly code was written to provide the output file attribute Breakoutimage which is available in DEC-10 SIMULA but unavailable in IBM SIMULA. This class appeared to solve many of the problems that were encountered.

The two versions of DIALOG are described in technical memoranda 79-3 and 79-3a and the former is no longer available because it is obsolete. As work with this class continued, two problems became quite obvious:

The declaration of class dialog is approximately 2500 lines of SIMULA and is supported by about 300 lines of assembly code. This imposes a significant overhead when compiling the verifier or the APL implementation. A single compile was costing about $20.

Both the APL verifier and the APL implementation must interact with terminals and files written in the standard character set (ASCII or EBCDIC) as well as the APL character set. There are three APL character sets: key-paired, bit-paired and tty-code. Only the first two are of interest for this work. However, character set translation was required so that the running program would be unaware of the character set of the input/output device. This translation is, quite properly, part of the input/output system and it was quite obvious that an extension of DIALOG to provide these capabilities would be both inefficient and quite unreliable because it would extend DIALOG far beyond what was anticipated in the original design.

Therefore, a comprehensive design of a file system that is a proper extension of the SIMULA Common Base Definition was undertaken. The view of files adopted in this design is quite different from the standard CMS view and is much easier to use.

The first assumption is that a file is named only by file specification. Moreover, simply writing to a file after giving its file specification is sufficient to cause the file to exist. Running programs may dynamically reference both input and output files using only the file specification. The concept of a DD name does not exist. If the specification of a file changes between executions of a program, then the file specification is read as data.

The second assumption is that a file may be associated with any input or output device and from the vantage point of a running program there is no difference between devices except for the text of the file specification. The various file formats of CMS are of no concern to the program.

The third assumption is that the data in a file may be written in one of three character sets: EBCDIC, key-paired APL or bit-paired APL. Further, a program may interpret characters coming from a file or written to a file as either a sequence of

EBCDIC characters or as a sequence of key-paired APL characters. It is the responsibility of the file system to provide appropriate translations in accord with the file specification.

A design requirement is that the files actually read and written must be standard CMS files so that programs written using this file system can communicate with other programs by way of the CMS file system.

A second design requirement is that all run time errors must be detected and reported to the user with an opportunity for corrective action or termination of execution at the option of the user.

The design of this file system is described in Technical Memorandum 79-8a and a part of the file system has been implemented. This design permits the incremental implementation of the file system and sufficient code to meet current needs has been written and tested. This code is being using in both the verifier and the APL implementation. As additional capabilities are required, they can be added to the system without modifying existing programs.

It has been decided to devote only the effort that is needed to provide a useful input/output system to this work and the present design makes it possible to defer the implementation of parts of the system until they are really needed.

It appears to be the case that the problem of adequately communicating with the user at a terminal and with disk files is under control and is no longer an obstacle to the main purpose of this grant.

In spite of rather limited distribution of technical memoranda, this work has attracted significant attention and the partial implementation of the input/output system is being used at other installations. The Norwegian Computing Center (IBM SIMULA implementors) is now referring users with input/output problems to the Principal Investigator and responding to inquiries consumes some time.

5. Other Software Development

The incremental verifier parses programs and assertions using an SLR(1) parser. The parser in the verifier reads the state table of the parser and a list of the tokens of the language from an input file and, therefore, it is independent of the source language of program to be verified.

These input files are produced from a BNF grammar by an SLR(1) parser generator that was written in DEC-10 RATFOR at the

University of Arizona by Dr. F. C. Druseikis. Both the RATFOR
and Fortran source files for this program were transferred to
Virginia Tech. This program consists of approximately 4000 lines
of RATFOR source code. The Fortran source code is substantially
unacceptable to the IBM Fortran compiler. In addition, the pro-
gram relies on the representation of integers on the host hard-
ware. To simplify the transfer of this program as well as some
small utilities and to provide a generally useful tool for pro-
gram development, a RATFOR progressor that is the same as the
UNIX and TOPS-10 version but which writes Fortran that is
accepted by the IBM compiler was constructed using part of the
code from this RATFOR implementation.

This RATFOR processor is now in general use at Virginia Tech
and has been exported to other CMS installations.

This RATFOR processor has been used for work on implementing
the SLR(1) parser generator on VM/CMS. Serious problems were
encountered when bringing this program into operation. The ori-
gin of these problems was character representation and integer
representation on the host computer system. The most important
parts of the parser generator are now operational but the report
printing program is not yet operational. This program prints
many of the error messages and other diagnostics that occur when
the input grammar is not appropriate and some effort is still
needed to bring the program into production.

A variety of utility programs to deal with problems that are
related to the physical location of the Principal Investigator
and to limitations in VM/CMS system utilities were constructed.

The Principal Investigator is located at the Virginia Tech
Northern Virginia Graduate Center at Dulles International Airport
and computing facilities are much more limited than they are on
the main campus. For example, the remote batch station line
printer is equipped with a BCD character set and is quite useless
for printing the text of SIMULA programs. Programs to simulate
an upper and lower case line printer on a DECwriter were written
as were programs to expand tabs in files before they are printed.

6. Publication Plans

Approximately one third of a monograph, "A Primitive Recur-
sive Semantics and Implementation of APL", has been completed.
This manuscript will be submitted to publishers in February 1980
and the entire manuscript should be completed by June 1980.

Extensive work on the development of an interactive file
system has lead to some insights into the necessary properties of
a file system in an interactive environment. A short paper on
this work will be submitted to Software Practice and Experience.

With a verifier in operation, a number of additional examples to substantiate the claim that this is a useful approach to verification will be completed in collaboration with Dr. D. E. Britton of RCA Laboratories. This work should result in a paper for publication in a computer science journal.

By the end of 1980, some significant APL programs should be verified and this should lead to another publication on the verification of APL programs.

Although informal contacts are not a substitute for direct publication, there is some evidence to suggest that the interactive environment that has been created for the work of this grant is generally useful in the VM/CMS environment. The fragment of the file system that currently exists has been expoted to Intel Corporation and is being used to solve similar problems here. The Norwegian Computing Center routinely refers other users with similar problems to the Principal Investigator.

The existence of CMS RATFOR is just becoming known and requests for copies of the system are being received. The University Computing Center has been asked to provide distribution services for this program and it is expected that this arrangement will be completed.

7. Personnel

Dr. Richard J. Orgass directed his efforts toward bringing the software that is used in this research into operation in the VM/CMS environment and toward completing the verification of the APL implementation discussed above.

To solve the file management problems, he designed and implemented the SIMULA class DIALOGG for the CMS environment and designed and partially implemented an interactive file system. This work on the file system is documented in technical memoranda.

He spent a significant amount of time working with the Norwegian Computing Center to solve problems related to bringing the verifier into operation and to constructing a usable working environent out of parts of the verifier and other software.

Dr. Orgass was assisted in this work by a number of student employees of the grant; their activities are summarized below.

Mr. Stephen M. Choquette implemented the CMS RATFOR system with general guidance from Dr. Orgass. Mr. Choquette received a B.S. in Computer Science from Virginia Tech in June 1979 and is now employed by IBM and a parttime graduate student in the Northern Virginia Graduate Program at Virginia Tech. He devoted 100 hours of his time to this effort.

Mr. Robert Porter was employed as a Graduate Research Assistant during the summer of 1979. He assisted Dr. Orgass in the implementation of DIALOG and supported the work of Dr. J. J. Martin (described below).

Mr. R. D. Johnson, a sophomore at Virginia Tech, has been working with Dr. Orgass on the implementation of the interactive file system. He created much of the assembly code that is under the SIMULA coding of the system. He is currently designing a file handler that will directly use the CMS output facilities in place of the OS simulation facilities that are provided with SIMULA. This work should result in considerably improved performance for the file system and programs that are used in the grant.

Mr. J. W. Stibbards, a senior at Virginia Tech, was briefly involved in the work of the grant but he found that he had too many other commitments and withdrew from this work. This was a wise decision because his education and student activities are much more important for him.

Mr. R. Critz, a Virginia Tech sophomore taking a year off from formal education is working for the grant as a programmer. He has brought the SLR(1) parser generator into operation and has written several assembly procedures to facilitate the management of the CMS environment. He serves the grant in the important capacity of program librarian and keeps the grant software in good order. He will continue to provide programing support until August 1980. Mr. Critz relieves the Principal Investigator of a substantial amount of routine work and makes essential contributions in this capacit. In addition to contributing to the work of the grant, Mr. Critz is also learning a great deal about computing and this contributes to his education.

Dr. J. J. Martin, Associate Professor of Computer Science at Virginia Tech, devoted one month to the work of the grant. During this time he familiarized himself with the verification techniques and programs that are used in this work. He has written a report describing many of the capabilities of the SLR(1) parser which will be useful as work continues. At this time it is not clear if Dr. Martin is sufficiently interested in this work to continue his participation in the work of the grant. A decision will be made before the summer of 1980. If Dr. Martin should terminate his affiliation with the grant, he will be replaced by another faculty member.

One of the chronic problems of this work has been finding graduate assistants to work with Dr. Orgass in Northern Virginia. Until December 1979, it was necessary for student assistants to travel to Dulles Airport to use Virginia Tech computing facilities. Since all Nothern Virginia Computer Science students have fulltime positions elsewhere, the need to travel to Dulles for access to computing facilities imposed unacceptable demands on limited time.

In December 1979 a leased 9600 bps telephone line from Blacksburg to Dulles was installed and statistical multiplexors are attached to this line. Using this facility, metropolitan Washington dial-up computer ports are provided and this new equipment makes grant research much more attractive. Mr. James Wolf, a Northern Virginia graduate student, has just joined the work of the grant and should make important contributions as he becomes more involved in this work.

8. Publications

The technical memoranda that were published in the course of the current year's work are listed on the following page. Copies are attached to the original of this report.

Department of Computer Science
VIRGINIA POLYTECHNIC INSTITUTE
AND STATE UNIVERSITY
Graduate Program in Northern Virginia
P. O. Box 17186
Washington, D.C. 20041

TECHNICAL MEMORANDA

79-1 R. J. Orgass. Concerning Classes Within Classes. January 15, 1979.

79-2 R. J. Orgass. Line Printer Spooling on an ASCII Terminal. April 12, 1979.

79-3a R. J. Orgass. A SIMULA Class for Writing Interactive Programs. October 14, 1979.

79-4 S. M. Choquette and R. J. Orgass. CMS RATFOR System Manual. July 1, 1979.

79-5 S. M. Choquette and R. J. Orgass. CMS RATFOR User's Manual. July 1, 1979.

79-6 R. J. Orgass. CMS Software Notebook (First Edition). July 31, 1979.

79-7 R. J. Orgass. Converting DEC-10 SIMULA Programs to CMS SIMULA. August 3, 1979.

79-8a R. J. Orgass. Design for a CMS SIMULA File System with 4 Character Sets. November 5, 1979.

80-1 J. J. Martin. A SIMULA Program for SLR(1) Parsing. January 1980.

80-2 R. J. Orgass. Interim Scientific Report, AFOSR Grant 79-0021. February 15, 1980.